Name_____ **Student #** _____

STAT 201
Midterm Examination

Richard Lockhart 21 October 2004

Instructions:

1. This is a closed book exam.
2. You may use a calculator (with no wireless communications ability).
3. You may bring one sheet of notes.
4. You may also bring the tear out sheet of tables and notes from the text.
5. Sometimes, to do the problem, you will need to make assumptions. You should be clear and explicit about what assumptions you need to make the technique you are using reasonable.
6. The exam is out of 30.
7. **DON'T PANIC.**

1. In a study of dietary fat intake 1000 father-son pairs were examined. (Both father and son were adults in all pairs.) For each person the percentage of dietary calories received in the form of fat is measured. The fathers received an average of 35% of their dietary calories in the form of fat with an SD of 6 percentage points. The sons received an average of 30% of their dietary calories in the form of fat with an SD of 8 percentage points. The correlation between father and son was r=0.4.

   (a) About what percentage of the fathers receive more than 40% of their daily caloric intake in the form of fat? Be clear about any assumption you must make to do the problem. (4 marks)

   (b) If a father receives 28% of his calories in the form of fat about what percentage should you predict for the son? (5 marks)

(c) Suppose we select 36 families at random from this group of 1000 for a more detailed dietary assessment. You may assume that they are selected with replacement so that the selections are independent. What is the chance that the 25 selected fathers have an average between 34 and 36% of their daily dietary calories in the form of fat. (4 marks)

(d) In 50 of the families the father received less than 15% of his daily caloric intake in the form of fat. If we eliminate this group of 50 father-son pairs from our study will the correlation coefficient go up or down; that is, is the correlation coefficient for the other 950 pairs more than 0.4, less than 0.4, or still about 0.4? Explain with a graph. (2 marks)

(e) Consider the families where the father receives about 28% of his calories in the form of fat. Approximately what would be the standard deviation of the sons' percentage of daily caloric intake in the form of fat in these families? (1 mark)

2. An executive of a large supermarket chain discovers that the correlation between the total amount of overtime worked by cashiers at a store and the total number of bad cheques accepted at a store is 0.7. He recommends that a ban be placed on overtime, arguing that cashiers at the end of a long day are less careful. What is wrong with this thinking; your answer should include an alternative explanation of the observed correlation. (2 marks)

3. A large bank has loans outstanding on 100,000 pieces of real estate. At the last audit the average assessed value of the pieces of real estate was $150,000. The bank suspects that recent economic events mean that the real estate values may have fallen suddenly. A simple random sample of 400 of the outstanding loans shows an average present value of $139,600 with an SD of $160,000.

   Looking more closely at the data collected the bank president goes through the files to find the assessed values of the 400 sampled pieces of real estate at the time of the last audit. The figures average $145,000 with an SD of $165,000. He discovers, however, that those properties valued at over $400,000 at the last audit have decreased in value $20,000 each on average while those valued at under $100,000 have not decreased in value at all on average. He develops the following explanation of this observation. Owners of expensive properties have had to sell them. They have then taken the proceeds of the sales and bought less expensive properties thereby keeping up the prices of these properties. Identify a pitfall in the executive's reasoning. (2 marks)

4. A forester lays out 10 small plots of land scatterd over a large area of recently logged forest land. In each plot 20 seedlings are planted. The forester returns a year later and counts the total number of the 200 seedlings which are still alive. S/he plans to treat the number alive as having a Binomial distribution. Is this wise? Why or why not? (2 marks)

5. A company manufactures steel rods. Each day it produces 50000 rods. The day's production is judged acceptable if fewer than 10% of the 50000 rods have a breaking strength less than some specified level. In order to check this a simple random sample of 100 rods is tested at the end of each day. Assuming that in fact the day's production is just barely acceptable–that is, that 5000 of the 50000 rods have too low a breaking strength–what is the chance that 14 or more of the 100 sampled rods fail the strength test? I am expecting an approximate, not exact, answer. (5 marks)

6. I toss a fair coin and throw a thumbtack. The coin is equally likely to land either heads (H) or tails (T). The thumbtack has chance 1/3 of landing point up (U); otherwise it lands tipped over (O). Make a list of all the outcomes in the sample space for this experiment and show the probability of each outcome. Explain the rules you used to get these probabilities. (3 marks)